

Efficient “on-the-fly” calculation of Raman spectra from *ab-initio* molecular dynamics: Application to hydrophobic/hydrophilic solutes in bulk water

Pouya Partovi-Azar

Department of Chemistry, University of Paderborn,
Warburger Str. 100, D-33098 Paderborn, Germany

Thomas D. Kühne

Department of Chemistry and Institute for Lightweight Design with Hybrid Systems,
University of Paderborn, Warburger Str. 100, D-33098 Paderborn, Germany*

We present a computational method to accurately calculate Raman spectra from first principles with an at least one order of magnitude higher efficiency. This scheme thus allows to routinely calculate finite-temperature Raman spectra “on-the-fly” by means of *ab-initio* molecular dynamics simulations. To demonstrate the predictive power of this approach we investigate the effect of hydrophobic and hydrophilic solutes in water solution on the infrared and Raman spectra.

I. INTRODUCTION

Ab initio simulations of vibrational spectra, where the electronic degrees of freedom are explicitly taken into account, often provides important insights into the structure and dynamics of complex systems. Therefore, computer simulations nowadays represent an invaluable tool to rationalize and complement experimental measurements. The key quantity to compute the vibrational spectrum of a system is its dipole moment, which is the tendency of charge distribution towards inhomogeneity. From this point of view, Maximally Localized Wannier Functions (MLWFs) are particularly useful, since they allow to partition the total electronic density into individual fragment contributions.¹ These Wannier functions are defined as

$$w_n(\mathbf{r} - \mathbf{R}) = \frac{V}{(2\pi)^3} \int_{BZ} d\mathbf{k} e^{-i\mathbf{k} \cdot \mathbf{R}} \sum_{m=1}^J U_{mn}^{(\mathbf{k})} \psi_{m\mathbf{k}}(\mathbf{r}), \quad (1)$$

where $\psi_{m\mathbf{k}}(\mathbf{r})$ are Bloch functions, \mathbf{R} is a Bravais lattice vector, and V is the real-space primitive cell volume, while the integral is computed over the whole Brillouin zone². The unitary $J \times J$ matrix $U_{mn}^{(\mathbf{k})}$ is periodic with respect to the wave-vector \mathbf{k} , while $\psi_{m\mathbf{k}}(\mathbf{r})$ are the eigenstates of a system as obtained by an electronic structure method, such as density functional theory (DFT)³. To compute MLWFs, the total spread functional

$$S = \sum_n S_n = \sim_n \left(\langle w_n | r^2 | w_n \rangle - \langle w_n | \mathbf{r} | w_n \rangle^2 \right), \quad (2)$$

is minimized by appropriately chosen unitary rotations $U_{mn}^{(\mathbf{k})}$.¹ Thereof, one can use the expectation value of the periodic position operator $\hat{\mathbf{r}}$ in the Wannier representation, in order to find the centers of the localized functions for arbitrary symmetries.^{4–6} As such, the polarization of the electronic charge in a crystal, which in general has a periodic continuous distribution, can be unambiguously partitioned into localized contributions.^{7,8} Having the centers and the associated spreads of the MLWFs and

considering each MLWF as a charge distribution in space, it is not only possible to calculate the molecular and/or total dipole moments of a system, but also the corresponding time-correlation functions by means of *ab-initio* molecular dynamics (AIMD) simulations. The temporal Fourier transform of the autocorrelation between the total dipole moments $\mathbf{M}(t)$ at different times, t , is proportional to the infrared (IR) absorptivity, $\alpha(\nu)$, which after employing the harmonic approximation, can be expressed as^{9–11}

$$\alpha(\nu)n(\nu) \propto \int_0^\infty dt e^{i2\pi\nu t} \langle \dot{\mathbf{M}}(0) \cdot \dot{\mathbf{M}}(t) \rangle_{\text{cl}}, \quad (3)$$

where ν and $n(\nu)$ are the frequency and the index of refraction, respectively, while $\langle \dots \rangle_{\text{cl}}$ denotes the ensemble-average in classical statistical mechanics. By applying a periodic electric field,^{12–14} usually using the Berry phase approach,^{15,16} it is possible to obtain the polarizability tensor \hat{A} via

$$A_{ij} = -\frac{\partial M_i(\mathbf{E})}{\partial E_j}; \quad i, j = \{x, y, z\}, \quad (4)$$

where E_j denotes the Cartesian component j of the applied electric field, while M_i is the i th component of the total dipole moment. However, calculating the derivatives numerically based on density functional perturbation theory (DFPT)^{17–19} or using higher-order finite difference (FD) methods²⁰ is computationally rather expensive. The mean polarizability $\bar{A} = 1/3 \text{Tr}[\hat{A}]$ is the quantity that is usually measured in experiment. With \bar{A} known, one can obtain the isotropic Raman spectrum through the calculation of the autocorrelation between the polarizabilities^{11,21}

$$\sigma(\nu) \propto \nu \int_0^\infty dt e^{i2\pi\nu t} \langle \bar{A}(0) \bar{A}(t) \rangle_{\text{cl}}. \quad (5)$$

In this paper we present a novel computational technique to efficiently calculate Raman spectra, where the polarizability of a system is represented as a sum over Wannier function polarizabilities denoted as a function of Wannier function volumes, and consequently by their spread.

II. WANNIER POLARIZABILITY METHOD

Specifically, it has been shown that the molecular polarizability change linearly with the volume of the electronic cloud around a molecule.²² Due to the fact that the electronic properties play the main role in quantifying the polarizability, it is reasonable to assume that the total isotropic polarizability of the system can be obtained as a sum over the Wannier function polarizabilities. The essentially same idea has been recently used to combine the quantum harmonic oscillator model with Wannier functions to calculate the non-local dynamic electron correlation due to oscillating charges assigned to MLWFs.²³ As such, the polarizability assigned to i th MLWF is given by

$$A_i = \beta S_i^3, \quad (6)$$

where S_i is the spread of the i th MLWF and β is a proportionality constant. Therefore, the mean polarizability of the system can be written as

$$\bar{A} = \frac{1}{3} \text{Tr}[\hat{A}] = \frac{1}{3} \sum_i^{N_{WF}} A_i = \frac{\beta}{3} \sum_i^{N_{WF}} S_i^3. \quad (7)$$

At variance to the FD technique, the numerical effort reduces from six to just one single-point calculation plus an additional MLWF computation. This is to say that our novel MLWF-based method, which hereafter we will refer to as Wannier polarizability (WP) method, is at least five times more efficient than the conventional FD approach. The speed-up with respect to DFPT is similar.²⁴

The parameter β is determined by minimizing the Mean Absolute Relative Error (MARE) of the mean polarizabilities with respect to reference calculations using the FD approach. All of our DFT calculations were conducted using the CP2K/QUICKSTEP code²⁵ in conjunction with a very accurate TZV2PX Gaussian basis set²⁶ and the Perdew-Burke-Ernzerhof exchange-correlation functional²⁷ plus a damped interaction potential to approximately account for long-range dispersion interactions.²⁸ The mean polarizabilities per molecule for various water clusters containing 2 to 13 molecules as obtained using the FD scheme are shown in Fig. 1. The structures of the water clusters at Hartree-Fock level of theory were taken from the Cambridge Cluster Database.²⁹ The obtained results are in very good agreement with previously calculated polarizabilities for the same systems at DFT level of theory.^{30,31}

The optimized value of β for our WP method at minimum MARE is $\beta = 0.90$. The corresponding isotropic polarizabilities of the water clusters are also displayed in Fig. 1. Even though the absolute polarizabilities of the FD and WP methods slightly differ for the smallest water clusters, the qualitative behavior with respect to cluster size is similar, which immediately suggest that the deduced Raman spectra most likely differ only in their absolute intensities. To assess the latter, we have compared

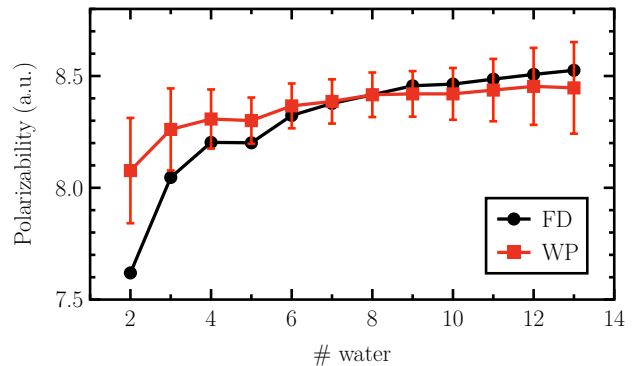


FIG. 1. Mean polarizabilities per water molecule for various water clusters containing 2 to 13 molecules as obtained using the FD (black circles) and WP (red squares) methods. The error bars for the red squares are the absolute relative errors at each point.

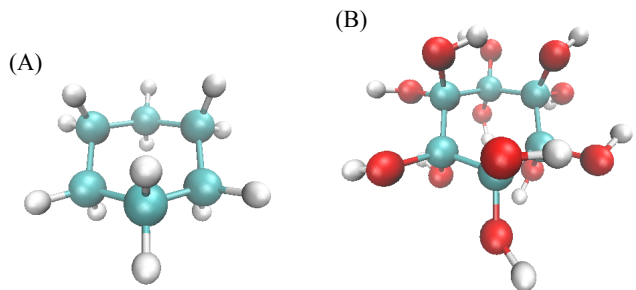


FIG. 2. Atomic configuration of cyclohexane (a) and cyclohexanodecol (b). Cyan, red and white spheres denote carbon, oxygen and hydrogen atoms, respectively.

the Raman spectra of a single cyclohexane molecule in the gas phase, shown in Fig. 2(a), using both FD and WP methods. The spectra were obtained “on-the-fly” from simulations using the second-generation Car-Parrinello method, where both the density matrix and $U_{mn}^{(k)}(t)$ were propagated together with the nuclei to further speed-up the calculations.^{32,33} However, for each AIMD step, the Wannier functions were re-localized using the scheme of Berghold et al. to obtain genuine MLWFs.⁶ Nonetheless, together with the WF method, this results in a combined acceleration for the calculation of the Raman spectra of at least an order of magnitude. At first the system was equilibrated in the canonical ensemble at 300 K for 10 ps using a discretized time-step of 0.5 fs, before sampling the polarizabilities in the micro-canonical ensemble for additional 10 ps. In the case of FD approach, at each time-step we applied an external electric field of 0.0001 a.u. intensity along the $\pm x$, $\pm y$ and $\pm z$ directions, to calculate the polarizabilities. The eventual Raman spectra are depicted in Fig. 3. As can be seen, the agreement between the FD and WP methods is excellent, and no frequency shift can be observed. The peaks around 800 cm^{-1} are typically attributed to C-C stretching and CH_2

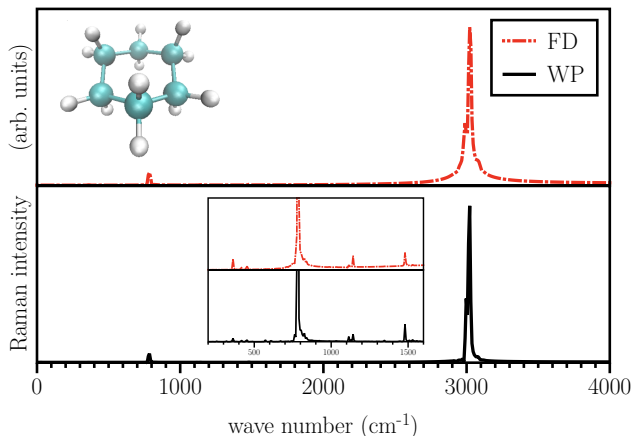


FIG. 3. Isotropic Raman spectra of cyclohexane molecule in the gas phase, as obtained by the FD (red dashed line) and WP (black straight line) methods.

rocking vibrations, while the peaks around 3000 cm^{-1} are usually assigned to symmetric and asymmetric CH_2 stretches. In the inset of Fig. 3 the Raman activity for frequencies less than 1500 cm^{-1} are shown, where the higher bands are due to other CH bending modes, while the lower ones originates from the torsion and deformation of the carbon ring^{34–41}. Our simulated spectra are in good agreement with other theoretical and experimental results.

III. RESULTS AND DISCUSSION

We demonstrate our novel WP method by investigating the effect of hydrophobic and hydrophilic molecules on the IR and Raman spectra of liquid water. To that extend we consider two systems in water solution: a single cyclohexane molecule (Fig. 2(a)) to represent a hydrophobic solute, and the relatively similar, but hydrophilic cyclohexanododecol molecule (Fig. 2(b)), where all hydrogen atoms are replaced by OH groups. Hereafter we refer to these two systems as CW and COHW, respectively. Both, CW and COHW contains 128 light water molecules per unit cell, where the water density is set to the experimental value at ambient conditions. Contrary to our previous calculations, a smaller DZVP basis set has been employed. Again, both structures were first equilibrated in the canonical ensemble at room temperature for $\sim 15\text{ ps}$, before the dipole moments and polarizabilities were sampled “on-the-fly” for 10 ps in the micro-canonical ensemble. The simulated IR and Raman spectra of CW and COHW systems are shown in Fig. 4. The IR activities for the frequencies below 650 cm^{-1} , at $\sim 1600\text{ cm}^{-1}$, as well as the weak peak at $\sim 2200\text{ cm}^{-1}$ are due to libration and hydrogen bond bending and stretching, O-H bending, and a combination of O-H bending and libration modes of bulk water,

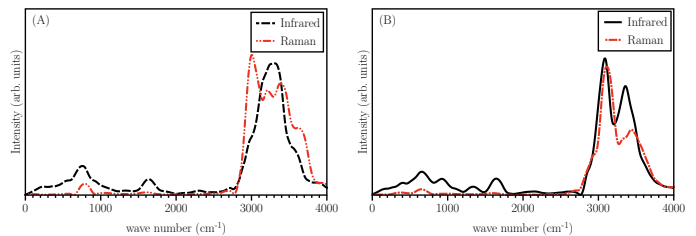


FIG. 4. IR and Raman spectra of CW (a) and COHW (b) systems. Black and red curves show infrared and Raman spectra, respectively.

respectively.^{42–46} The peak at $\sim 800\text{ cm}^{-1}$ in the IR and Raman spectra of the CW and COHW systems originate most likely from C-C stretching modes of the carbon ring.^{34–41} Moreover, IR activities in the frequency range $950\text{--}1100\text{ cm}^{-1}$ can be assigned to C-O stretching modes of cyclohexanododecol, while the peak at $\sim 1350\text{ cm}^{-1}$ seems to be due to its C-O-H bending mode.⁴⁷ By comparison of Fig.4(a) with Fig.4(b) we attribute the IR-active peak at $\sim 3000\text{ cm}^{-1}$ of the CW system to C-H stretching modes of the cyclohexane molecule. The remaining frequencies above 3000 cm^{-1} are the O-H stretching modes.^{48,49} However, at variance to bulk water, we observe a generally larger splitting between the symmetric and asymmetric stretching modes, which immediately suggests that the asymmetry of the hydrogen-bond network is more pronounced due to the presence of the solute.^{49–51} In the case of the hydrophobic cyclohexane molecule the effect is even more pronounced. Moreover, for the CW system the so-called dangling O-H bond peak at $\sim 3650\text{ cm}^{-1}$ is more distinct and gives rise to a shoulder as can be seen in Fig. 4(a).^{52–55} We believe that the latter is a consequence of fleetingly broken hydrogen-bonds of the distorted hydrogen-bond network which is spanned around the hydrophobic solute.⁵⁶

IV. CONCLUSIONS

In summary, we have presented a novel method that allows to efficiently calculate IR and in particular Raman spectra “on-the-fly” within AIMD simulations. To that extend we exploit the fact MLWFs, which are at the core of this new approach, can be utilized to partition the charge distribution of the system into localized fragments. Therefore, the total isotropic polarizability can be calculated as a sum over the Wannier polarizabilities, which are assumed to be proportional to its volume and determined by its spread. Together with an extension of the second-generation Car-Parrinello method to propagate $U_{mn}^{(k)}(t)$ along with the nuclei, followed by a re-localization to obtain genuine MLWFs, a speed-up of one order of magnitude has been observed. Using this approach, we calculated IR and Raman spectra for cyclohexane and cyclohexanododecol solutes in ambient bulk

water. We found that the former hydrophobic solute give rise to a shoulder at around $\sim 3650\text{ cm}^{-1}$, which is due to momentarily dangling O-H bonds. In any case, we

conclude by noting that this development facilitates to routinely calculate finite temperature spectra with only minimal extra computational cost.

* tdkuehne@mail.uni-paderborn.de

- ¹ N. Marzari and D. Vanderbilt, *Physical Review B* **56**, 12847 (1997).
- ² G. H. Wannier, *Physical Review* **52**, 191 (1937).
- ³ R. O. Jones and O. Gunnarsson, *Rev. Mod. Phys.* **61**, 689 (1989).
- ⁴ P. L. Silvestrelli, N. Marzari, D. Vanderbilt, and M. Parrinello, *Solid State Commun.* **107**, 7 (1998).
- ⁵ P. L. Silvestrelli, *Physical Review B* **59**, 9703 (1999).
- ⁶ G. Berghold, C. J. Mundy, A. H. Romero, J. Hutter, and M. Parrinello, *Physical Review B* **61**, 10040 (2000).
- ⁷ R. Resta, *Physical Review Letters* **80**, 1800 (1998).
- ⁸ R. Resta and D. Vanderbilt, in *Physics of Ferroelectrics* (Springer, 2007) pp. 31–68.
- ⁹ R. Ramírez, P. Kumar, and D. Marx, *The Journal of chemical physics* **121**, 3973 (2004).
- ¹⁰ R. Iftimie, P. Minari, and M. E. Tuckerman, *Proceedings of the National Academy of Sciences of the United States of America* **102**, 6654 (2005).
- ¹¹ P. Atkins and J. De Paula, *Atkins' physical chemistry* (Oxford University Press, 2010).
- ¹² P. Umari and A. Pasquarello, *Physical Review Letters* **89**, 157602 (2002).
- ¹³ I. Souza, J. Íñiguez, and D. Vanderbilt, *Physical Review Letters* **89**, 117602 (2002).
- ¹⁴ P. Umari and A. Pasquarello, *Physical Review B* **68**, 085114 (2003).
- ¹⁵ R. D. King-Smith and D. Vanderbilt, *Physical Review B* **47**, 1651(R) (1993).
- ¹⁶ R. Resta, *Rev. Mod. Phys.* **66**, 899 (1994).
- ¹⁷ P. Giannozzi and S. Baroni, *The Journal of chemical physics* **100**, 8537 (1994).
- ¹⁸ S. Baroni, S. De Gironcoli, A. Dal Corso, and P. Giannozzi, *Reviews of Modern Physics* **73**, 515 (2001).
- ¹⁹ A. Putrino and M. Parrinello, *Phys. Rev. Lett.* **88**, 176401 (2002).
- ²⁰ I. Vasiliev, S. Ögüt, and J. R. Chelikowsky, *Physical review letters* **78**, 4805 (1997).
- ²¹ T. Ishiyama, V. V. Sokolov, and A. Morita, *The Journal of chemical physics* **134**, 024510 (2011).
- ²² J. S. Murray, P. Lane, T. Brinck, K. Paulsen, M. E. Grice, and P. Politzer, *The Journal of Physical Chemistry* **97**, 9369 (1993).
- ²³ P. L. Silvestrelli, *The Journal of Chemical Physics* **139**, 054106 (2013).
- ²⁴ A. Putrino, D. Sebastiani, and M. Parrinello, *The Journal of Chemical Physics* **113**, 7102 (2000).
- ²⁵ J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter, *Computer Physics Communications* **167**, 103 (2005).
- ²⁶ J. VandeVondele and J. Hutter, *The Journal of chemical physics* **127**, 114105 (2007).
- ²⁷ J. P. Perdew, K. Burke, and M. Ernzerhof, *Physical review letters* **77**, 3865 (1996).
- ²⁸ S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, *The Journal of chemical physics* **132**, 154104 (2010).
- ²⁹ S. Maheshwary, N. Patel, N. Sathyamurthy, A. D. Kulkarini, and S. R. Gadre, *The Journal of Physical Chemistry A* **105**, 10525 (2001).
- ³⁰ M. Yang, P. Senet, and C. Van Alsenoy, *International journal of quantum chemistry* **101**, 535 (2005).
- ³¹ T. K. Ghanty and S. K. Ghosh, *The Journal of chemical physics* **118**, 8547 (2003).
- ³² T. D. Kühne, M. Krack, F. R. Mohamed, and M. Parrinello, *Physical Review Letters* **98**, 066401 (2007).
- ³³ T. D. Kühne, *WIREs Comput. Mol. Sci.* **4**, 391 (2014).
- ³⁴ M. Ito, *Spectrochimica Acta* **21**, 2063 (1965).
- ³⁵ K. B. Wiberg and A. Shrake, *Spectrochimica Acta Part A: Molecular Spectroscopy* **27**, 1139 (1971).
- ³⁶ E. Mátrai, P. Császár, G. Fogarasi, and M. Gál, *Spectrochimica Acta Part A: Molecular Spectroscopy* **41**, 425 (1985).
- ³⁷ T. Hirschfeld and B. Chase, *Applied spectroscopy* **40**, 133 (1986).
- ³⁸ S. Frankland and M. Maroncelli, *The Journal of chemical physics* **110**, 1687 (1999).
- ³⁹ M. Pelletier, *Applied spectroscopy* **53**, 1087 (1999).
- ⁴⁰ B. Jordanov, D. Tsankov, and E. Korte, *Journal of molecular structure* **651**, 101 (2003).
- ⁴¹ P. Kukura, D. W. McCamant, and R. A. Mathies, *Annu. Rev. Phys. Chem.* **58**, 461 (2007).
- ⁴² Y. Maréchal, *The Journal of chemical physics* **95**, 5565 (1991).
- ⁴³ D. M. Carey and G. M. Korenowski, *The Journal of chemical physics* **108**, 2669 (1998).
- ⁴⁴ P. Larouche, J.-J. Max, and C. Chapados, *The Journal of chemical physics* **129**, 064503 (2008).
- ⁴⁵ J.-J. Max and C. Chapados, *The Journal of chemical physics* **131**, 184505 (2009).
- ⁴⁶ Q. Sun, *Chemical Physics Letters* **568**, 90 (2013).
- ⁴⁷ J. Antony, G. von Helden, G. Meijer, and B. Schmidt, *Journal of Chemical Physics* **123**, 014305 (2005).
- ⁴⁸ T. D. Kühne, M. Krack, and M. Parrinello, *Journal of Chemical Theory and Computation* **5**, 235 (2009).
- ⁴⁹ C. Zhang, R. Z. Khaliullin, D. Bovi, L. Guidoni, and T. D. Kühne, *The Journal of Physical Chemistry Letters* **4**, 3245 (2013).
- ⁵⁰ T. D. Kühne and R. Z. Khaliullin, *Nature communications* **4**, 1450 (2013).
- ⁵¹ T. D. Kühne and R. Z. Khaliullin, *Journal of the American Chemical Society* **136**, 3395 (2014).
- ⁵² Q. Du, R. Superfine, E. Freysz, and Y. R. Shen, *Phys. Rev. Lett.* **70**, 2313 (1993).
- ⁵³ Q. Du, E. Freysz, and Y. R. Shen, *Science* **264**, 826 (1994).
- ⁵⁴ N. Ji, V. Ostroverkhov, C. S. Tian, and Y. R. Shen, *Physical Review Letters* **100**, 096102 (2008).
- ⁵⁵ S. Nihonyanagi, T. Ishiyama, T.-k. Lee, S. Yamaguchi, M. Bonn, A. Morita, and T. Tahara, *Journal of the American Chemical Society* **133**, 16875 (2011).
- ⁵⁶ D. Chandler, *Nature* **437**, 640 (2005).